Supplementary Materials for

Chromatin state dynamics during blood formation

David Lara-Astiaso, Assaf Weiner, Erika Lorenzo-Vivas, Irina Zaretsky, Diego Adhemar Jaitin, Eyal David, Hadas Keren-Shaul, Alexander Mildner, Deborah Winter, Steffen Jung, Nir Friedman,* Ido Amit*

*Corresponding author. E-mail: nir@cs.huji.ac.il (N.F.); ido.amit@weizmann.ac.il (I.A.)

Published 7 August 2014 on *Science* Express
DOI: 10.1126/science.1256271

This PDF file includes:

iChIP

iChIP is a modular protocol (**Fig. 1A**) in which every step has been optimized to efficiently handle low cell amounts. 1) Cells are cross-linked sorted and frozen in aliquots of 10-20 thousand cells; 2) chromatin of 500-10,000 cells is sheared in a low volume sonication and 3) immobilized on antibody coated magnetic beads. 4) Bead immobilized chromatin is indexed by ligation of sequencing adaptors. 5) The indexed chromatin is released from the antibody coated magnetic beads using antibody denaturing conditions and pooled together with other samples in a single tube, 6) the chromatin pool is washed to remove antibody denaturing elements. 7) The chromatin pool is subjected to a chromatin immunoprecipitation step by adding magnetic beads coated with the target ChIP antibody, 8) Bead bound immunocomplexes are washed to remove non-specific binding, RNA and proteins are degraded and DNA is reverse crosslinked. 9) ChIPed DNA, which contains P5 and P7 Illumina sequences, is purified using SPRI AMPure XP beads. 10) DNA is amplified by PCR. Table S1 summarizes the primers used in iChIP.

1. Cell Harvesting

FACS-Sorted cross-linked cells are collected in 5 ml FACS tubes containing 500 ul of fetal calf serum, diluted in 4 ml of iChIP Harvesting Buffer (12mM TrisCl, 0.1X PBS, 6 mM EDTA, 1.2X Protease Inhibitors (Roche) and pelleted by centrifugation for 15 min at 1000xG using a swing rotor with low acceleration and brake settings. Supernatant is removed leaving a 200 ul cover to avoid disturbing the cell pellet. Cells are re-suspended in the 200 ul buffer cover, transferred to 0.2ml tubes in aliquots of 10,000 -20,000 cells, then cells are pelleted at 1000xG for 15 min using above settings. Supernatant is removed leaving 10 ul of buffer covering the cell pellet. Cells are frozen stored at -80 C.

2. Sonication

Cell aliquots are thawed on ice and 2 ul of 3%SDS is added to achieve a final concentration of 0.5% SDS. Cells are thoroughly re-suspended and lysed on ice for 10 min. Lysates were transferred into 0.1 ml Bioruptor Microtubes (Diagenode C30010015, Liège, Belgium) and the chromatin was sheared using an NGS Bioruptor Sonicator (Diagenode) at High Intensity and cycles of 30" ON/30" OFF. Sonication time was calibrated for the different cell types: 20 min for MEPs, EryA and EryB; 25 min for CMPs; 30 min for LT-HSC, ST-HSC, MPP and GMPs; 40 min for Granulocytes, Monocytes and Macrophages and 45 min for NK, T and B cells.

3. Chromatin Immobilization

After sonication, chromatin extracts were transferred to a 96 well plate and diluted 1 to 5 with Sonication Equilibration Buffer (10 mM TrisCl, 140 mM NaCl, 0.1 % Sodium Deoxycholate, 1% Tx-100, 1 mM EDTA, 1X Protease Inhibitors (Roche) to achieve an SDS concentration of 0.1%. To immobilize the chromatin on magnetic beads, 15 ul of Dynabeads Protein (Life Tech) and 1.3 ug of anti-H3 antibody (ab1791) were added to the diluted chromatin extracts and incubated for 20 hours at 4 degrees.

4. Chromatin Indexing

Magnet based bead capture was used to efficiently add, wash and remove the different master mixes used in the indexing process. All the reactions were done while chromatin was bound to the H3 coated magnetic beads. After the immobilization IP bead bound chromatin immunocomplexes were magnetized and washed 3 times with 150 ul of 10 mM Tris Cl + 1X Protease Inhibitors EDTA free (Roche). After the washes bead bound chromatin was re-

suspended in 20 ul of the same buffer. Chromatin End Repair was performed by adding 30 µl of a master mix: 25 µl 2X ER mix (50 mM Tris-Cl ph 7.5, 20 mM $MgCl_2$, 20 mM DTT, 2 mM ATP, 1 mM dNTPs), 2 ul T4 PNK enzyme (10 U/ul NEB), and 2 µl T4 polymerase (3 U/ul NEB) to each well. Samples were incubated in a thermal cycler at 12C for 25 min, 25C for 25 min, and finally cooled to 4C. After end repair, bead bound chromatin was washed once with 150 ul of 10mM TrisCl + Protease Inhibitors and re-suspended in 40 ul of the same buffer. Chromatin was A-tailed by adding 20 µl master mix (17 µl A-base add mix, 3 µl Klenow (3'->5' exonuclease, 3 U/ul, NEB) to each well and incubated at 37C for 30 min. in a thermal cycler. After end repair bead bound chromatin was washed once with 150 ul of 10mM TrisCl + Protease Inhibitors and re-suspended in 19 ul of the same buffer. Chromatin was indexed by adding 5 ul of 0.75 uM Y-shapped Indexed Adaptors (containing P5 and P7 sequences) to each well which were ligated to the chromatin's DNA ends by adding 34 µl of AL master mix (29 µl 2x Quick Ligation Buffer and 5 µl Quick DNA ligase (NEB) to each well. Samples were thoroughly mixed and incubated at 25C for 40 min in a thermal cycler. Bead bound indexed chromatin was washed once with 150 ul of 10mM TrisCl + Protease Inhibitors to remove the non ligated adaptors.

5. Chromatin Release

Denaturing conditions (DTT, high salt and detergent) and heat were used to release the indexed chromatin from the antibody coated magnetic beads. Right after the post-Indexing wash, samples were taken out of the magnet, beads were re-suspended in 12.5 ul of 100 mM DTT and incubated for 5 min at Room Temp. Then, 12.5 ul of Chromatin Release Buffer (500mM NaCl, 2% SDS, 2% Sodium Deoxycholate, 2X protease Inhibitors) were added to each well, samples were mixed thoroughly and incubated at 37C for 30 min. After the release incubation, magnetic beads were again thoroughly re-suspended and pooled together in groups of n samples resulting in a pool volume of 200-250 ul. The pooled indexed chromatin samples were concentrated using a 50Kda cutoff Centricon (Amicon).

6. Chromatin Immunoprecipitation

Target antibody was added and incubated at 4C for 3h, then 50 ul with Protein G Magentic beads were added and IP was incubated for 1 more hour. For each ChIP, we used 1.5 ug of anti H3K4me1 (ab8895) and 2.5 ug of anti H3K4me2 (ab32356), anti H3K4me3 (Millipore, 07-473) and anti H3K27ac (ab4729). TF iChIP: 50 ul of protein G magnetic Dynabeads (Invitrogen) coupled to target antibody were added and incubated 4h at 4 degrees. For the coupling, beads were washed once (200ul) in a binding/blocking buffer (PBS, 0.5% Tween 20, 0.5% BSA), incubated with 10 ug of antibody in binding/blocking buffer for 1 hour at room temperature, and then washed to remove excess antibody. For each ChIP, we used 10 ug of anti PU.1 (Santa cruz, sc352-x).

7. Washes and ChIPed DNA elution

A 96 well magnet was used (Invitrogen) in all further steps. ChIP Buffer was removed and samples were washed 5 times with cold RIPA (200ul per wash), twice with RIPA buffer supplemented with 500 mM NaCl (200ul per wash), twice with LiCl buffer (10 mM TE, 250mM LiCl, 0.5% NP-40, 0.5% DOC), once with TE (10Mm Tris-HCl pH 8.0, 1mM EDTA), and then eluted in 50 ul of 0.5% SDS, 300 mM NaCl, 5 mM EDTA, 10 mM Tris HCl pH 8.0. The eluate was treated sequentially with 2ul of RNaseA (Roche, 11119915001) for 30 min at 37 C, 2.5 ul of Proteinase K (NEB, P8102S) for two hours at 37 C and 8 hours at 65C to revert formaldehyde crosslinking.

8. ChIPed DNA isolation.

SPRI cleanup steps were performed using 96 well plates and magnets. 90ul SPRI were added to the reverse-crosslinked samples, pipette-mixed 15 times and incubated for 6 minutes. Supernatant

were separated from the beads using a 96-well magnet for 5 minutes. Beads were washed on the magnet with 70% ethanol and then air dried for 5 minutes. The DNA was eluted in 23 ul EB buffer (10 mM Tris-HCl pH 8.0) by pipette mixing 25 times.

9. Library Amplification, QC and Sequencing.
The library was completed and amplified through a PCR reaction with 0.5 µM of PCR forward and PCR reverse primers and PCR ready mix (Kapa Biosystems). The forward primer contains the Illumina P5-Read1 sequences and the reverse primer contains the P7-Read2 sequences (see table S7). The amplified pooled single-cell library was purified with 1x volumes of SPRI beads. Library concentration was measured with a Qubit fluorometer (Life Technologies) and mean molecule size was determined with a 2200 TapeStation instrument (Agilent Technologies). Library quality was further determined by qPCR, measuring the enrichment of ubiquitous active promoters (Actin-B and GAPDH) versus background (Cryaa). Libraries were size selected with 0.6x volumes of SPRI beads removing large DNA fragments (>600 bp). iChIP libraries were sequenced using an Illumina HiSeq 1500.

RNA isolation
A minimum of 5000 cells were sorted in 200 ul of Lysis/Binding Buffer (Life technologies), lysed for 5 min and frozen at -80C. Cell lysates were thawed and messenger RNA was captured with 12 ul of Dynabeads oligo(dT) (Life technologies), and washed according to manufacture guidelines. Purified messenger RNA was eluted at 70C with 10 ul of 10 mM Tris-Cl ph 7.5 and stored at -80C.

RNAseq library construction
We used the MARS-seq protocol (*29*) developed for single cell RNA-seq to produce RNA-seq libraries for all the hematopoietic populations studied in this article. In brief, the protocol consists of special designed primers with unique molecular identifiers for accurate molecule counting and a step of linear amplification of the initial mRNA pool, followed by a library construction step. This way, the diversity of the original pool of messenger RNAs is preserved even if the amount of input RNA is low. A minimum of 2 replicate libraries was prepared for each of the different hematopoietic populations. Table S1 summarizes the primers used in MARS-Seq.

The RNAseq protcol used for hematopoietic cells is described below:

1. Linear Amplification of the mRNA pool
4 ul of purified mRNA were placed in 384-well plates. First, to open secondary RNA structures and allow annealing of the RT primer, the 384-well plate was incubated at 72˚C for 3 min and immediately transferred to 384-well Inheco thermal block integrated to Bravo and set at 4˚C. Then, 2 ul of an RT reaction mix (10 mM DTT, 4 mM dNTP, 2.5 U/µl Superscript III RT enzyme in 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$) were added into each well of the 384-well plate and the reaction was mixed one time. Tips were replaced and the process repeated to all wells. The 384-well plate was then spun down and moved into a 384 cycler (Eppendorf) for the following incubation: 2 min at 42˚C, 50 min at 50˚C, 5 min at 85˚C. Indexed samples with equivalent amount of cDNA were pooled. _The pooled cDNA was converted to double-stranded DNA with a second strand synthesis kit (NEB) in a 20µl reaction, incubating for 2.5 h at 16˚C._The product was purified with 1.4x volumes of SPRI beads, eluted in 8 µl and in-vitro transcribed (with the beads) at 37˚C overnight for linear amplification using the T7 High Yield RNA polymerase IVT kit (NEB). Following IVT, the DNA template was removed with Turbo DNase I (Ambion) 15 min at 37˚C and the amplified RNA (aRNA) purified with 1.2x volumes of SPRI beads.

2. Library preparation for high-throughput sequencing
The aRNA was chemically fragmented into short molecules (median size ~200 nucleotides) by incubating 3 min at 70˚C in Zn^{2+} RNA fragmentation solution (Ambion) and purified with two volumes of SPRI beads. The aRNA (5 µl) was preincubated 3 min at 70˚C with 1 µl of 100 µM ligation adapter; then, 14 µl of a mix containing 9.5% DMSO, 1 mM ATP, 20% PEG8000 and 1 U/µl T4 ligase in 50 mM Tris HCl pH7.5, 10 mM $MgCl_2$ and 1mM DTT was added. The reaction was incubated at 22˚C for 2 h._The ligated product was reverse transcribed using Affinity Script RT enzyme (Agilent; reaction mix contains Affinity Script RT buffer, 10 mM DTT, 4 mM dNTP, 2.5 U/µl RT enzyme) and a primer complementary to the ligated adapter. The reaction was incubated for 2 min at 42˚C, 45 min at 50˚C and 5 min at 85˚C. The cDNA was purified with 1.5x volumes of SPRI beads._The library was completed and amplified through a nested PCR reaction with 0.5 µM of P5_Rd1 and P7_Rd2 primers and PCR ready mix (Kapa Biosystems). The forward primer contains the Illumina P5-Read1 sequences and the reverse primer contains the P7-Read2 sequences. The amplified pooled library was purified with 0.7x volumes of SPRI beads to remove primer leftovers. Library concentration was measured with a Qubit fluorometer (Life Technologies) and mean molecule size was determined with a 2200 TapeStation instrument (Agilent). MARS-Seq libraries were sequenced using an Illumina HiSeq 1500.

ATAC-seq
To profile for open chromatin, we used the Assay for Transposase Accessible Chromatin (ATAC-seq) protocol developed by Buenrostro et al. (*23*) with the following changes: different hematopoietic cell populations were sorted in 400ul of MACS buffer (1x PBS, 0.5% BSA, 2mM EDTA) and pelleted by centrifugation for 15min at 500g and 4ºC using a swing rotor with low acceleration and brake settings. Cell pellets were washed once with 1x PBS and cells were pelleted by centrifugation using the previous settings. Cell pellets were re-suspended in 25ul of lysis buffer (10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.1% Igepal CA-630) and nuclei were pelleted by centrifugation for 30min at 500g, 4C using a swing rotor with low acceleration and brake settings. Supernatant was discarded and nuclei were re-suspended in 25 ul reaction buffer containing 2ul of Tn5 transposase and 12.5ul of TD buffer (Nextera Sample preparation kit from Illumina). The reaction was incubated at 37ºC for one hour. Then 5ul of clean up buffer (900mM NaCl, 300mM EDTA), 2ul of 5% SDS and 2ul of Proteinase K (NEB) were added and incubated for 30min at 40C. Tagmentated DNA was isolated using 2x SPRI beads cleanup. For library amplification, two sequential 9-cycle PCR were performed in order to enrich small tagmentated DNA fragments. We used 2ul of indexing primers included in the Nextera Index kit and KAPA HiFi HotStart ready mix. After the first PCR, the libraries were selected for small fragments (less than 600 bp) using SPRI cleanup. Then a second PCR was performed with the same conditions in order to obtain the final library. DNA concentration was measured with a Qubit fluorometer (Life Technologies) and library sizes were determined using TapeStation (Agilent Technologies). Libraries where sequenced on a Hiseq 1500 for an average of 20 million reads per sample

***Ex vivo* differentiation of BMDCs.**
To obtain BMDCs, bone marrow cells were plated at a density of 200,000 cells/ml on non-tissue culture treated plastic dishes (10ml medium per plate). At day 2, cells were fed with another 10ml medium per dish. At day 5, cells were harvested from 15ml of the supernatant by spinning at 1400rpm for 5 minutes; pellets were resuspended with 5ml medium and added back to the original dish. Cells were fed with another 5ml medium at day 7. BMDC medium contains: RPMI (Gibco) supplemented with 10% heat inactivated FBS (Gibco), β-mercaptoethanol (50uM, Gibco),

L-glutamine (2mM, Biological Industries) penicillin/streptomycin (100U/ml, Biological Industries), MEM non-essential amino acids (1X, Biological Industries), HEPES (10mM, Biological Industries), sodium pyruvate (1mM, Biological Industries), and GM-CSF (20 ng/ml; Peprotech).

Harvesting of BMDCs for iChIP titration
Cells were fixed for 8 min with 1% formaldehyde, quenched with glycine and washed twice with ice-cold PBS. Cells were re-suspended in iChIP Harvesting Buffer, counted and transferred to 0.2 ml tubes in 10 ul aliquots containing 10.000, 3.000, 1.000 and 500 cells.

Isolation of hematopoietic progenitor cells
Femora, pelvis and tibiae were extracted from 6 C57BL/6J female mice (8 to 12 weeks old) and bone marrow cells were flushed with MACS buffer. The cells were enriched with CD117/c-kit Microbeads according to the manufacturer's guidelines (AutoMACS, Miltenyi Biotec; order number: 130-091-224). C-kit enriched cells were stained with Lin (Ter119, Gr1, CD11b, B220, CD3, CD4, CD8), c-kit, Sca1, CD34, FcgR-II, Flk2 and IL7R antibodies (eBioscience) for 30 min and sorted with FACSAria III cell sorter (BD Biosciences). Table S2 summarizes the antibodies used for this experiment. The cell populations were identified as:
LT-HSC: Lin-, c-Kit+, Sca-1+, Flk2-, CD34-
ST-HSC: Lin-, c-Kit+, Sca-1+, Flk2-, CD34+
MPP: Lin-, c-Kit+, Sca-1+, Flk2+, CD34+
CMP: Lin-, c-Kit+, Sca-1+, FcgRII low, CD34+
GMP: Lin-, c-Kit+, Sca-1+, FcgRII high, CD34+
MEP: Lin-, c-Kit+, Sca-1+, FcgRII-, CD34-
CLPs: Lin-, Flk2+, Il7R+

Isolation of bone marrow derived mature cells
Bone marrow cells were flushed from femora and tibiae of C57BL/6J female mice (8 to 12 weeks old), suspended with MACS buffer and incubated 3 min in red blood cell lysis solution (Sigma). Cells were stained with CD3, B220, NK1.1, Gr1, CD11b, CD115, F4/80 antibodies (eBioscience), and after filtration through a 70-µm strainer sorted with a FACSAria III cell sorter (BD Biosciences). Table S2 summarizes the antibodies used for this experiment. The cell populations were identified as:
Granulocytes: CD3-, B220-, Nk1.1-, CD11b+, Gr1+, High SSC
Macrophages: CD3-, B220-, F4/80+, CD115-, Low SSC
Monocytes: CD3-, B220-, NK1.1-, F4/80-, CD115+ Low SSC

Isolation of splenic lymphoid cells
Spleens were extracted from C57BL/6J female mice (8 to 12 weeks old), dissociated with a gentle MACS Dissociator (Miltenyi Biotec, Germany) into single suspension, and - after washing with MACS buffer - incubated for 5 min in red blood cell lysis solution (Sigma). Cells were then washed, resuspended in MACS buffer and stained with: CD3, CD4, CD8, B220, CD19, TCR-β, I-Ab, Ter119 and NK1.1 (eBioscience). After washing, cells were filtrated through a 70-µm strainer and sorted with a FACSAria III cell sorter (BD Biosciences). Table S2 summarizes the antibodies used for this experiment. The cell populations were identified as:
B cells: CD3-, B220+, CD19+.
T CD4+ cells: CD3+, B220-, CD4+, CD8-.
T CD8+ cells: CD3+, B220-, CD4-, CD8+.
NK Cells: CD4-, CD8-, B220-, I-Ab-, Ter119-, TCR-β-, NK1.1+.

Isolation of splenic erythroid cells

Spleens were extracted from C57BL/6J female mice (8 to 12 weeks old), dissociated into single splenocytes in MACS buffer, with a gentle MACS Dissociator (Miltenyi Biotec, Germany) and stained with: B220, Ter119 and CD71 (eBioscience). After washing, cells were filtrated through a 70-µm strainer and sorted with a FACSAria III cell sorter (BD Biosciences). Table S2 summarizes the antibodies used for this experiment. The cell populations were identified as:
EryA: Ter119+, CD71+, high FSC.
EryB: Ter119+, CD71+, low FSC.

ChIP sequencing and peak finding
ChIP sequencing was done on Illumina HiSeq-1500 and pooled libraries were sequenced at a sequencing depth of ~10-15 million aligned reads per sample. Libraries were prepared in triplicates or duplicates. Reads were mapped to the mouse mm9 assembly using the 'bowtie2' program with the default parameters, only tags that uniquely mapped to the genome were used for further analysis. Due to the initial low cell number per sample (~10,000) and to avoid clonal artifacts introduced in the PCR amplification step, we restricted the number of identical reads at each genomic position to be at most 3.

The identification of ChIP-seq enriched regions (peaks) in each sample was performed using HOMER (http://homer.salk.edu/homer/). For histone modifications, peaks were identified by searching locations of high read density using a 1000 bp sliding window. We required adjacent peaks to be at least 1000 bp away to avoid redundant detection. The threshold for the number of tags that determined a valid peak was selected at a false discovery rate of 0.001. The following HOMER command was used:

```
cmd = findPeaks <sample tag directory> -L 0 -C 3 -size 1000 -
minDist 1000 -tbp 3 -o <output file>
```

Construction of enhancer catalog
Peaks from all samples (16 cell types with 2-3 replicates) were combined into one unified catalog for each modification separately. Peaks that did not overlap between at least two replicates (distance between centers > 500bp) were discarded; appearance in one cell type was enough to be included in the final catalog. We removed redundancy of overlapping peaks (distance between centers < 500bp) by selecting a representative peak as the one with the strongest signal, extended by 1000bp in each direction from its center. Scatter plots and correlations between replicates were calculated on the vector of log-transformed reads counts for the collection of peaks (**Figs. 1D**, **S1A**).

In total, the catalog of H3K4me1 and H3K4me2 had 110,844 and 69,970 peaks, respectively, with 66,338 common peaks. To generate the final enhancer catalog, we first counted the number of H3K4me3 reads in each putative peak. We observed a bi-modal distribution of H3K4me3 levels and fitted a 2-Gaussian mixture model to select a threshold (**fig. S13A**). We next removed regions with high H3K4me3 levels: resulting with 48,415 regions in the final set of enhancers.

From this point onward, we used this catalog of enhancers for the rest of the analysis in this work. For each enhancer in this catalog, we counted the number of reads within 2kb around its center for each modification. We used the script annotatePeaks.pl from HOMER package, normalizing each library to 10,000,000 reads.

Clustering and correlation analysis
The enhancer catalog was clustered based on H3K4me1 log-transformed counts using K-means (matlab R2012b kmeans implementation) with K=9. We have tested other k (5-20) and observed

that increasing the number of clusters leads to quantitative refinements of the patterns, but does not introduce new patterns of enhancer change. Correlation matrices for histone modifications between cell types (fig S2, S3) were calculated on our log transformed enhancers catalog.

Enhancer dynamics and state

To define the number of dynamic enhancers, we first divided enhancers into 3 categories for each cell type: 'off' (number of reads < 25), 'on' (number of reads > 50), and 'intermediate' otherwise (**fig. S13B**). Enhancers that were categorized at least once in both 'on' and 'off' state were classified as dynamic. To quantify the number of enhancers that are opened or closed during the differentiation process, we counted the number of 'off'/'on' enhancers at the root of the hematopoietic tree (LT-HSC) that flipped state in the relevant mature cell type. To further determine the precise stage of gain/loss (**Figs. 3D, S6C**), we examined the path from HSC to mature cells and selected the stage in which the maximal change in H3K4me1 levels occurred.

Enhancer activity was determined based on the ratio of H3K4me1 and H3K27ac levels. We focused only on the 'on' regions based on H3K4me1 when defining the 'active' / 'poised' enhancers. As opposed to the bi-modal distribution observed in H3K4me1 levels over the set of regions, H3K27ac showed a normal distribution. Therefore, we selected a threshold for 'active' regions of H3K27ac reads > 25. Changing the threshold influenced the 'active'/'poised' ratio but did not change the relative order of cell types. For example, erythrocytes utilize more enhancers than CMP regardless of the specific threshold selected (**fig. S13C)**.

Analysis of ATAC-seq

ATAC-seq data was mapped and analyzed using the same pipeline as ChIP-seq data. We extracted a read coverage profile for each region in the enhancers catalog. The chromatin accessibility value for each enhancer in a given cell type was defined based on the maximum coverage in the region. We next re-centered our catalog based on the position of the maximum value in the ATAC coverage profile.

Motif enrichment analysis

Motifs were searched on 1000bp window centered on the ATAC center using homer2 command from the HOMER package (http://homer.salk.edu/homer/). For each region, we extracted the best motif score ('homer2 find –mscore') for each motif in the JASPAR CORE motif catalog (downloaded on March 2014)

http://jaspar.genereg.net//html/DOWNLOAD/JASPAR_CORE/pfm/nonredundant/pfm_vertebrates.txt

Enrichment scores were calculated using two-sample Kolmogorov-Smirnov test, comparing the motif scores in the set of 'off' enhancers with those in the 'on' enhancers for each cell type. Motifs corresponding to TFs not expressed in a certain cell type were excluded from the analysis.

Logistic model

To predict enhancer state ('on' or 'off') from its sequence, we trained a logistic model (*30*).
First, we estimated whether each transcription factor was expressed in each cell type. We used RNA-seq measurements, defining "expressed" as having more than 5 reads per million tags, and higher than 20% of its maximal expression levels across all cells.
The training labels were the state of each enhancer in each cell type, removing ones in intermediate state. The predictor was given the scores of the 216 motifs of the enhancer masked by the expression status of the corresponding transcription factor. Thus, the

prediction could utilize the motif scores of factors that are expressed in the relevant cell type. Parameters were estimated using matlab R2012b *glmfit* function. To evaluate accuracy, sensitivity and precision of the model we used 5-fold cross-validation, each time training on 80% of the enhancers X cell types and testing on the remaining 20%. The reported results are the average on 1000 cross validation runs, each with a random partition to 5 folds.

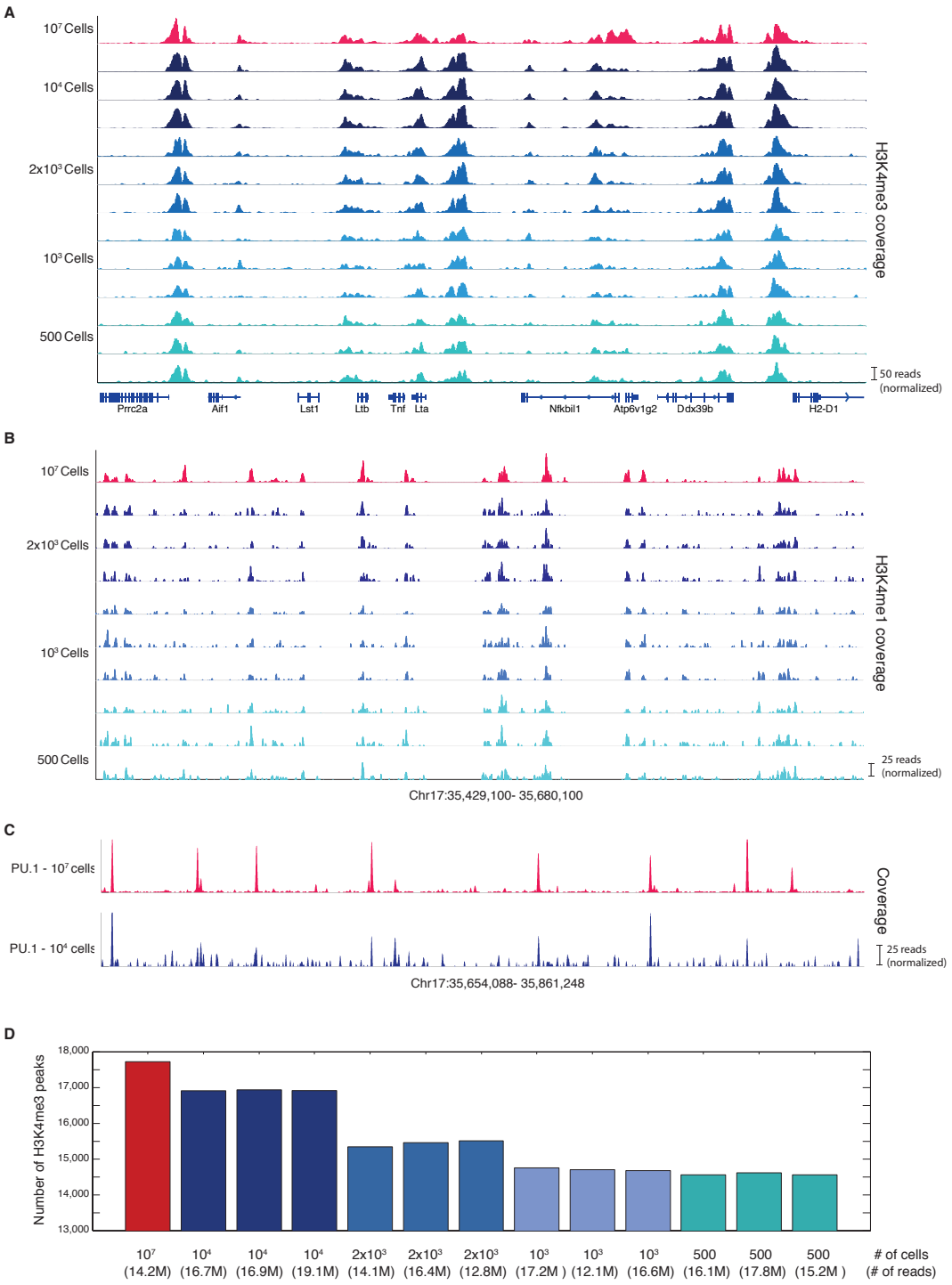


Figure S1. Indexing-first Chromatin ImmunoPrecipitation for profiling histone modifications and TF binding. **(A)** Normalized H3K4me3 profiles of peaks found in a 100 Kb region in the TNF locus (genes are indicated below) obtained with iChIP of decreasing amounts of bone marrow derived dendritic cells. **Top**, in red, H3K4me3 profile obtained using conventional ChIP with 20 million cells (*15*). Below, in shades of blue, replicates of H3K4me3

profiles. **(B)** Normalized H3K4me1 profiles of peaks found in a 250 Kb region in the TNF locus (coordinates are indicated below) obtained with iChIP of decreasing amounts of bone-marrow derived dendritic cells (BMDC). **Top**, in red, H3K4me1 profile obtained using conventional ChIP with 20 million cells (*15*). Below, in shades of blue, replicates of H3K4me1 profiles obtained with iChIP. **(C)** Normalized Pu.1 profiles of peaks found in a 250 Kb region in the TNF locus **Top**, in red, Pu.1 profile obtained using conventional ChIP with 20 million cells (*15*). Below, in blue, Pu.1 profile obtained with iChIP with 10^4 bone-marrow derived dendritic cells (BMDC). **(D)** The number of H3K4me3 enriched regions found at different libraries from **Fig S1A.**

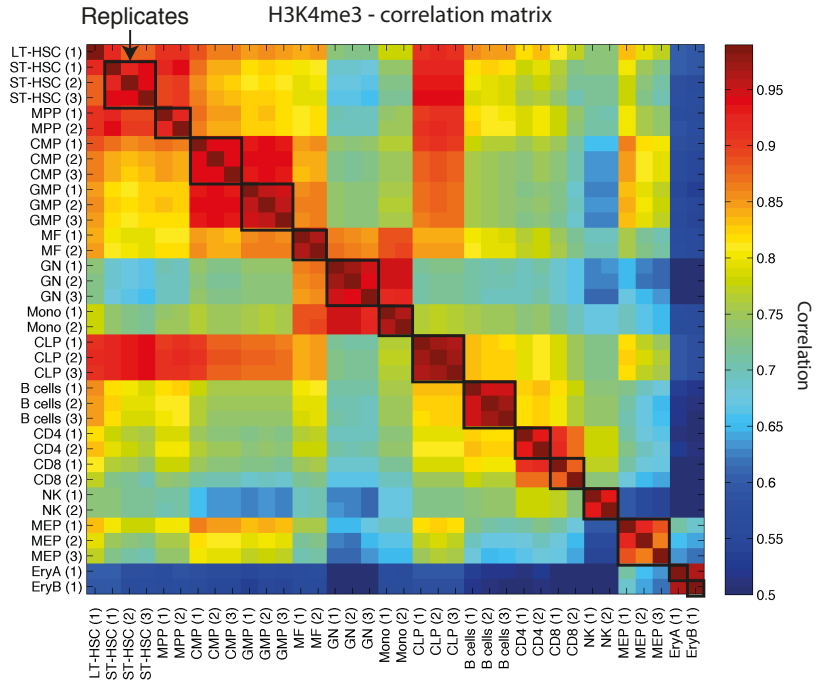


Figure S2. iChip is reproducible across biological replicate. Heatmap showing correlation of H3K4me3 peaks between iChIP replicates for all 16 cell types analyzed in this study.

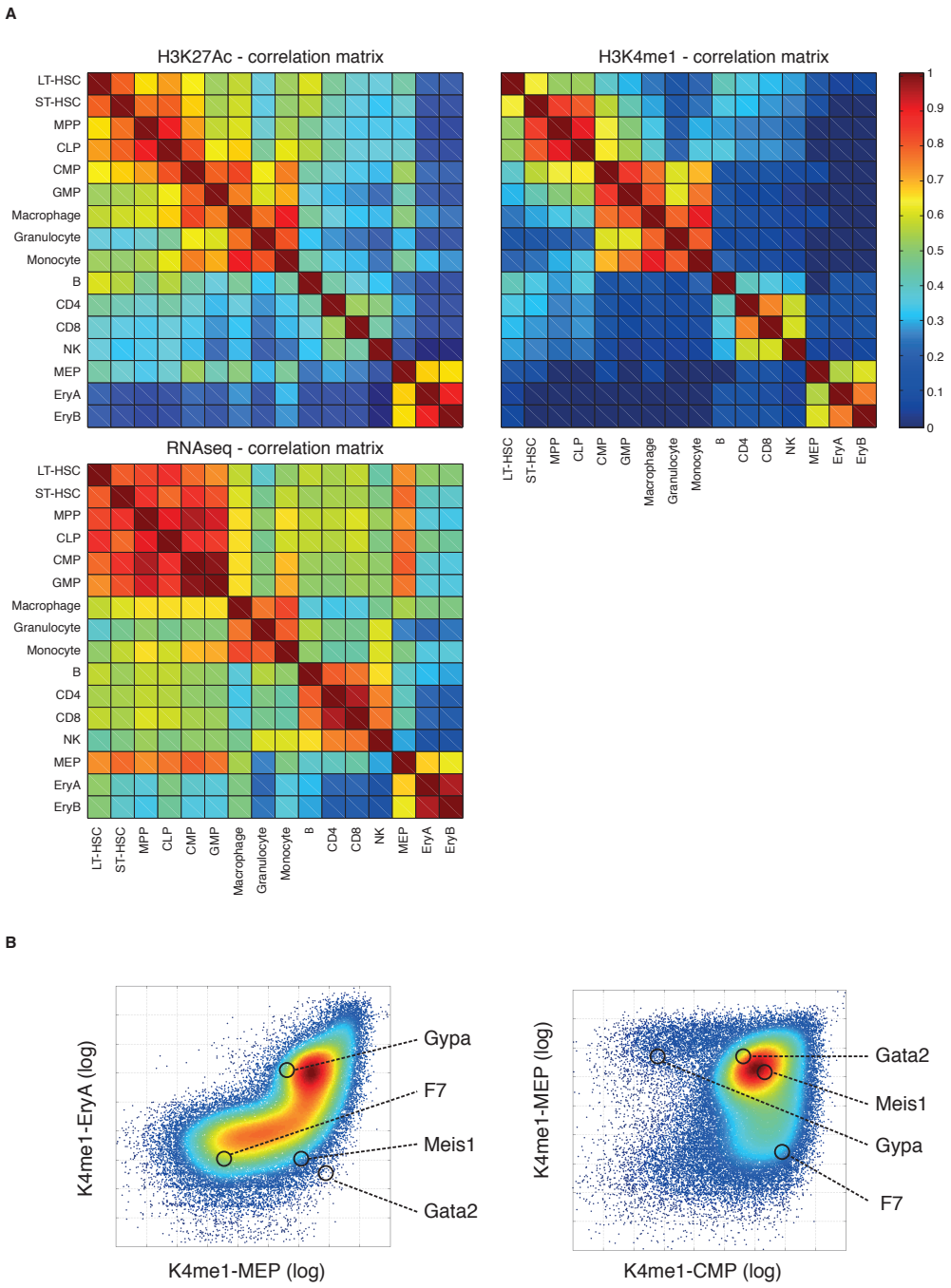


Figure S3. H3K4me1 patterns in progenitors display lineage potential prior to the execution of the gene expression program (A) Heatmap showing correlation of H3K27ac levels (top left) H3K4me1 levels (top right) and RNA expression levels for all the 16 cell types analyzed in this study. **(B)** Scatter plots of H3K4me1 read counts (log) in all 48,415 enhancer regions identified; upper panel compares MEP to its successive developmental step EryA. MEP to EryA transition (left) involves enhancer loss. Highlighted are enhancers in Gata2 and Meis1 loci. CMP and its descendant MEP (right), displaying both enhancer gain (e.g. Gypa) and loss (e.g. F7).

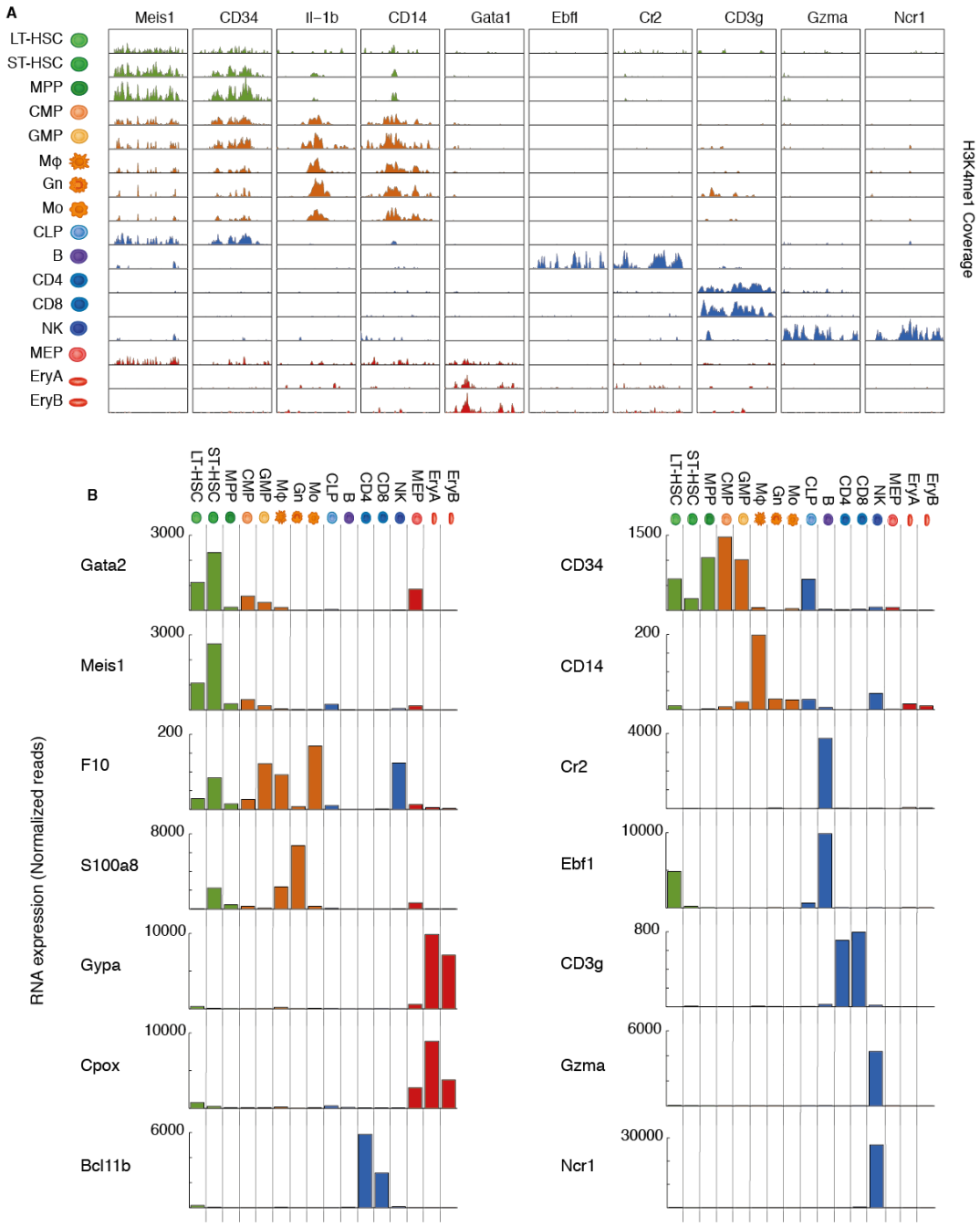


Figure S4. Chromatin and RNA dynamics in hematopoiesis (A) Representative examples of H3K4me1 signal (cell types are shown on the left) in several loci, (from left to right): CD34 and Meis1 for progenitors, IL-1b and CD14 for myeloid lineage, Gata1 for erythroid lineage, Ebf1 and Cr2 for B cells, CD3g for T cells, Gzma and Ncr1 for NK cells. Displayed are normalized read coverage in a variable region around the gene body. **(B)** Bar-plots showing gene expression profiles across the hematopoietic cell types for genes present in loci displaying lineage specific H3K4me1 signal (from **Figs. 2, S4A**).

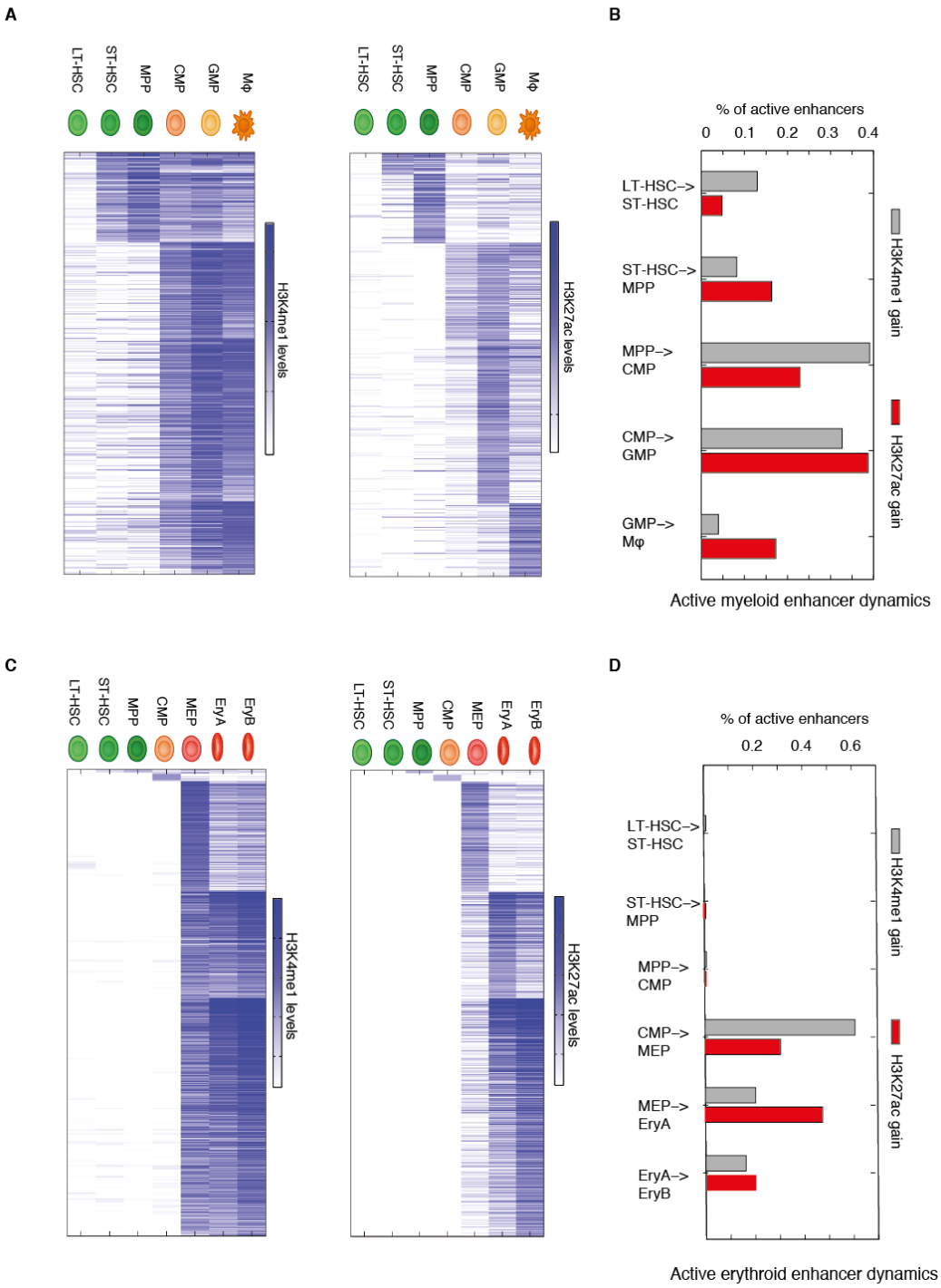


Figure S5. Stepwise acquisition of histone marks Heatmap showing H3K4me1 (left) and H3K27ac (right) signal over 1,556 and 3,104 active enhancers (H3K4me1/2 and H3K27ac) along macrophage **(A)** and erythroid (**C**) differentiation pathway respectively. The order of enhancers is sorted based on H3K27ac gaining step. Bar chart showing the percentage of enhancers that gain H3K4me1 or H3K27ac in each of the transitions from LT-HSC to macrophages (**B**) or EryB (**D**).

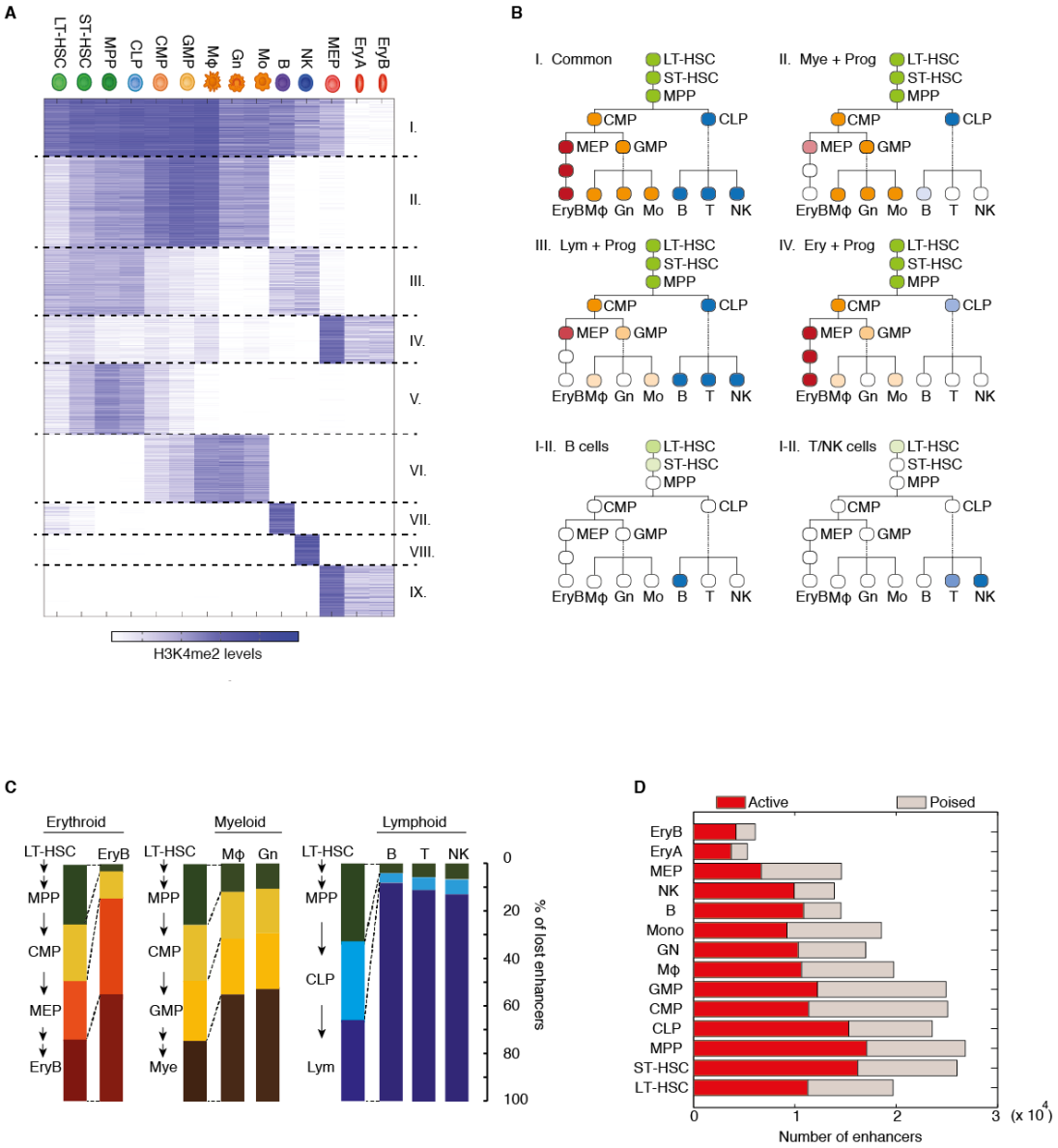


Figure S6. Global analysis of enhancer dynamics during hematopoiesis (A) Heatmap showing H3K4me2 signal in the 48,415 H3K4me1 enhancer regions used throughout hematopoiesis. The order of the enhancers is the same as in **Figure 3A. (B)** Schematic tree-view of the remaining six enhancer clusters; common (I), myeloid + progenitors (II), lymphoid + progenitors (III), erythroid + progenitors (IV), B cells (VII) and T/NK cells (VIII). **(C)** Bar plot showing the % of enhancers (from Fig. 3C) lost at each developmental stage from the HSC to mature hematopoietic cells. **(D)** Number of active (red) and poised (light gray) enhancers in each of the cell types studied.

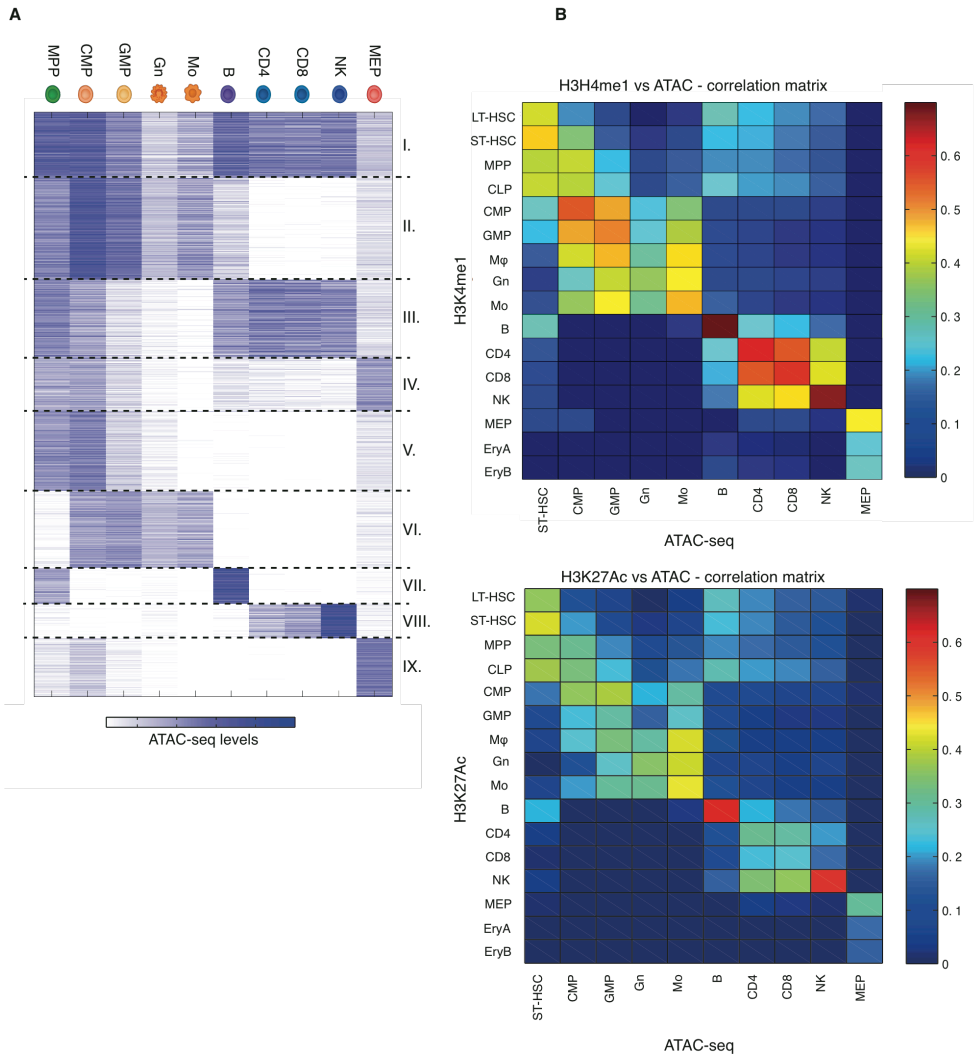


Figure S7. ATAC-seq dynamics in enhancers co-occur with establishment of H3K4me1 signal. (A) Heatmap showing ATACseq signal in the 48,415 enhancers used throughout hematopoiesis. The order of the enhancers is the same as in **Figure 3A**. **(B)** Heatmap showing correlation between ATAC-seq and H3K4me1 (top) or H3K27ac (bottom) for all 16 cell types analyzed (iChIP) versus all 10 cell types with ATAC-seq profiles.

Figure S8. **Cell stage specific motif enrichment**. Heatmap (same as **Fig. 4C)** showing the p-values of transcription factor motif enrichment (Kolmagorov-Smirnov test) for the indicated cell-type-specific enhancers (#). Red indicates significant motif enrichment (p< 1e-5), green indicates motif depletion (p<1e-5). White indicates either no significant motif enrichment or no RNA expression of the relevant transcription factor.

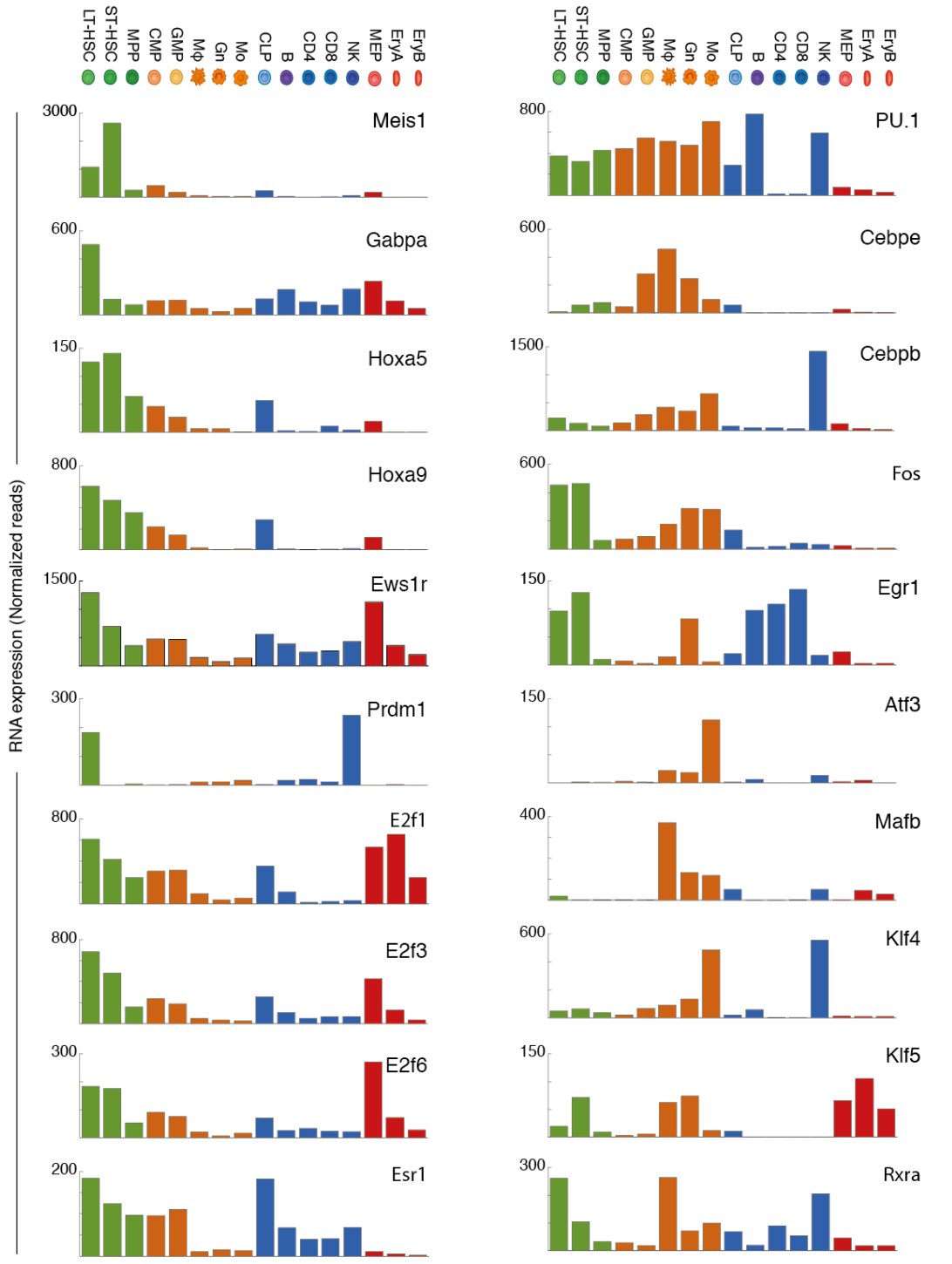


Figure S9 .Expression of lineage determining transcription factors identified as regulators of chromatin dynamics in hematopoiesis. Bar plots showing gene expression profiles across the hematopoietic cell types for representative transcription factors (from **Fig. 4C**) with putative lineage specific enhancer regulation.

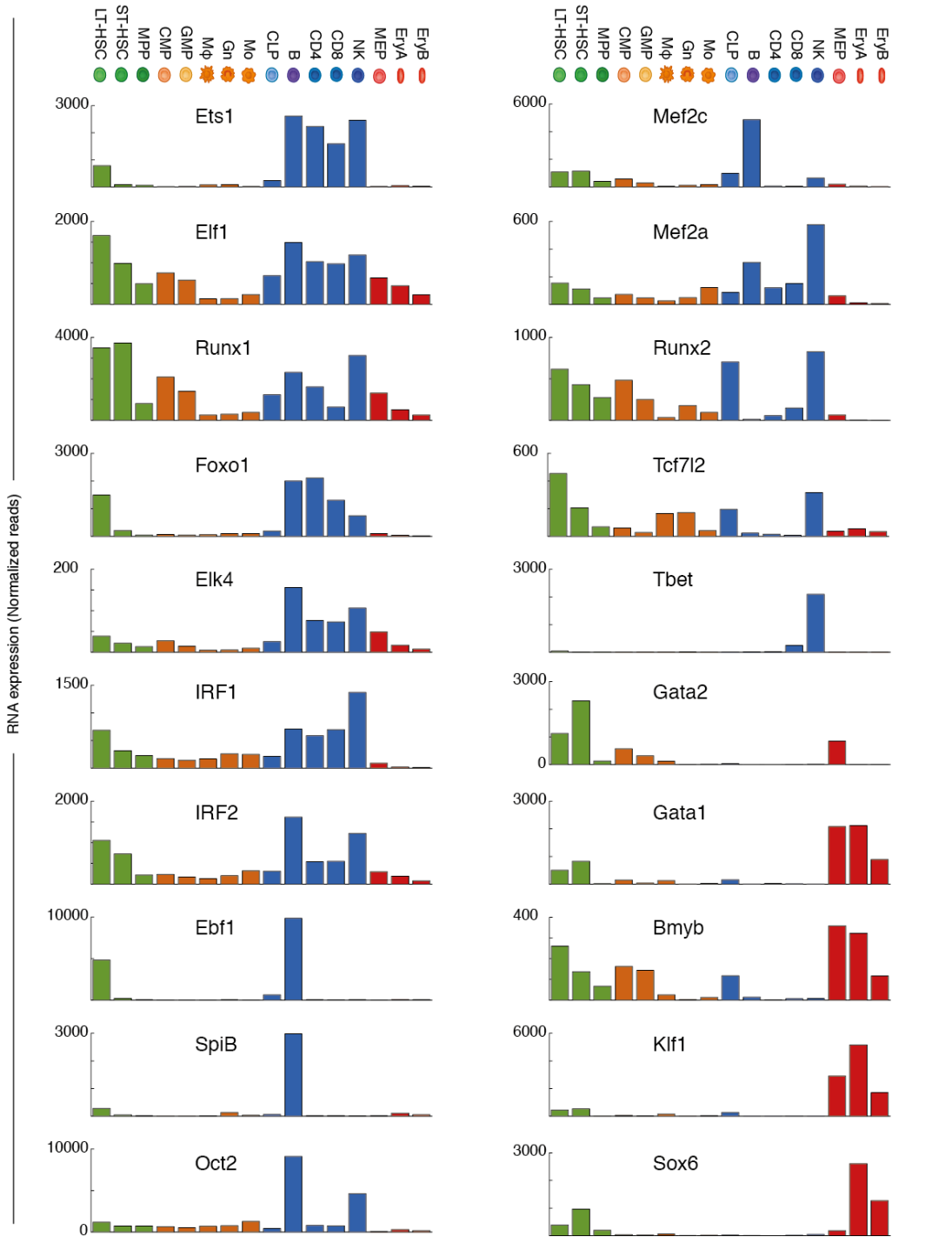


Figure S10. Expression of lineage determining transcription factors identified as regulators of chromatin dynamics in hematopoiesis. Bar plots showing gene expression profiles across the hematopoietic cell types for representative transcription factors (from **Fig. 4C**) with putative lineage specific enhancer regulation.

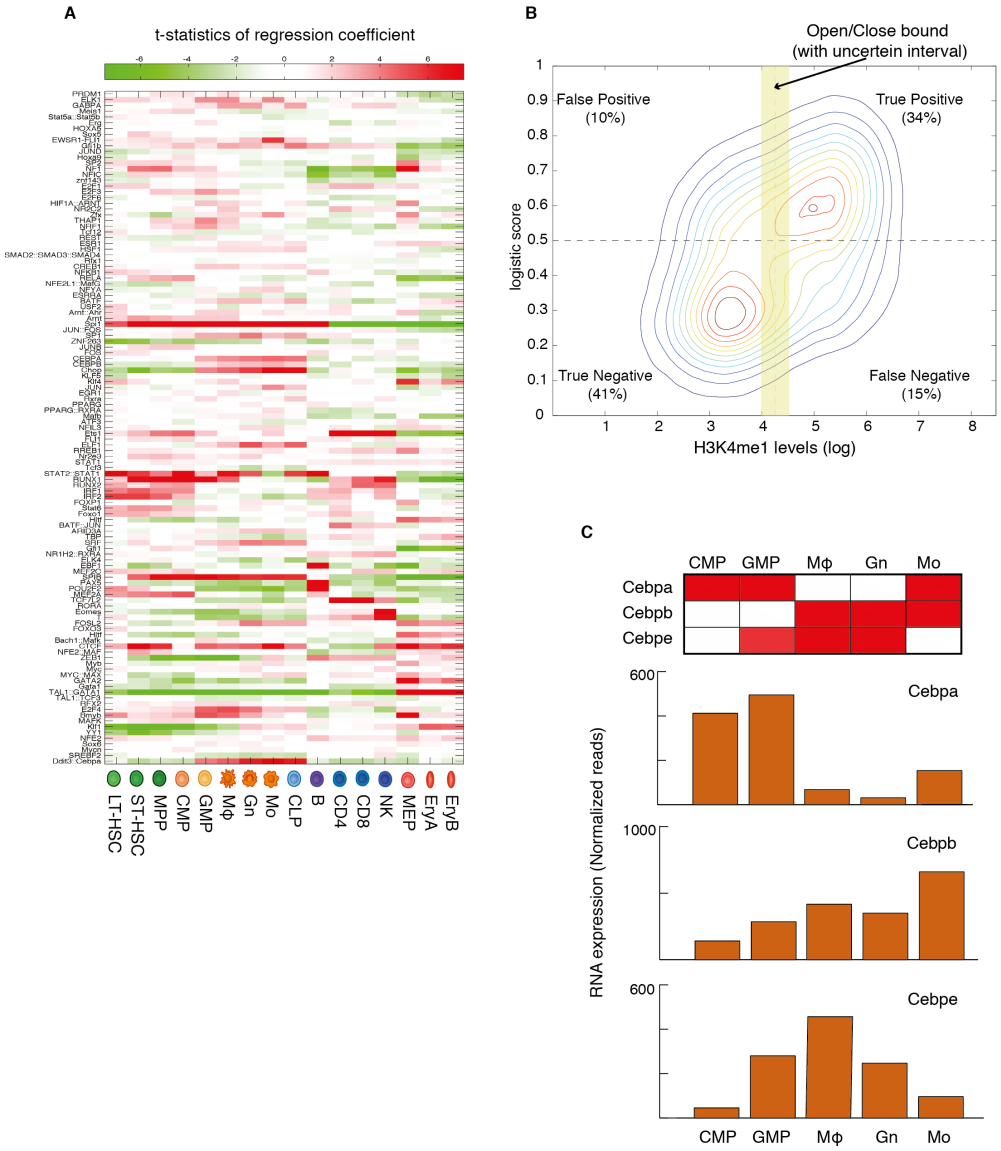
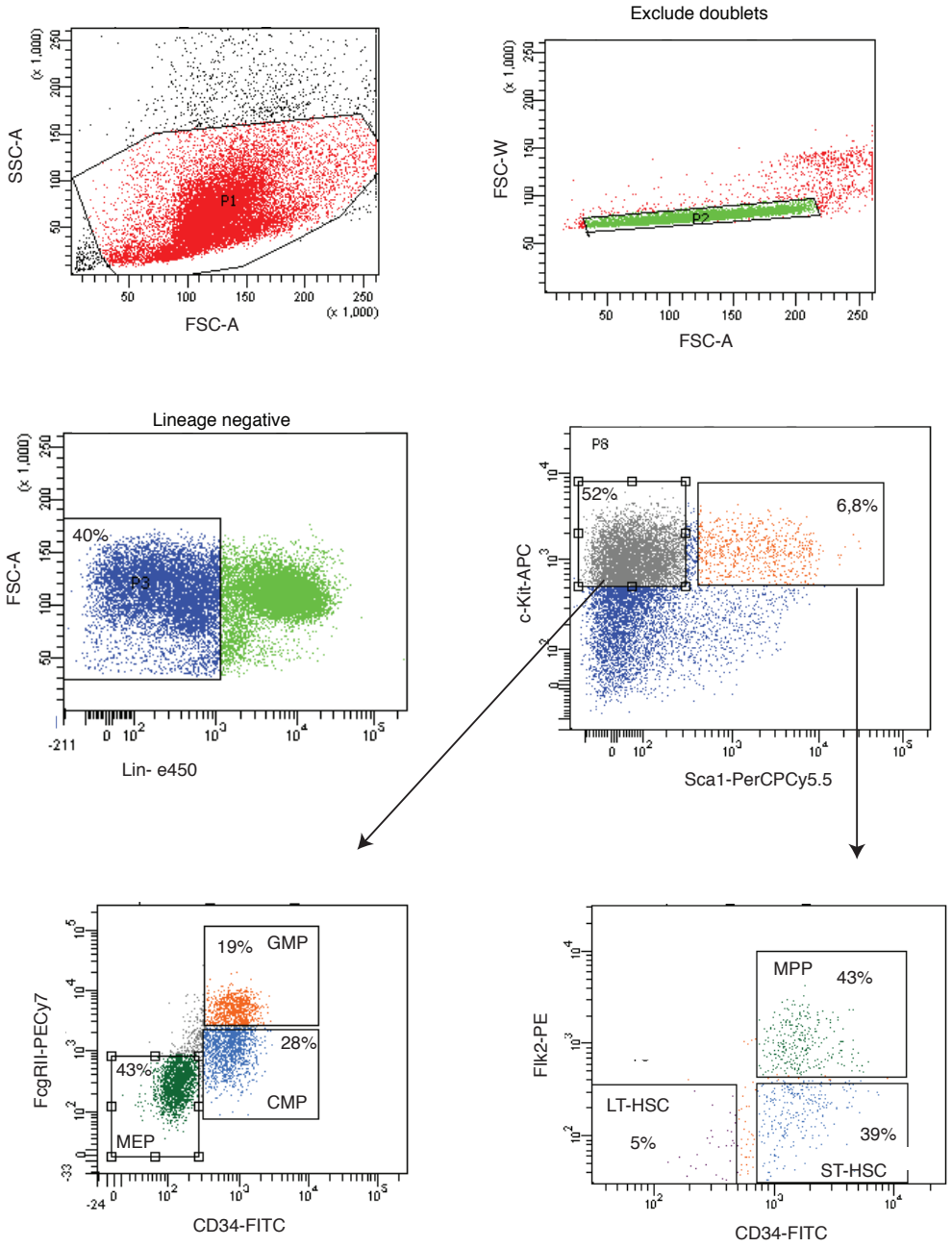
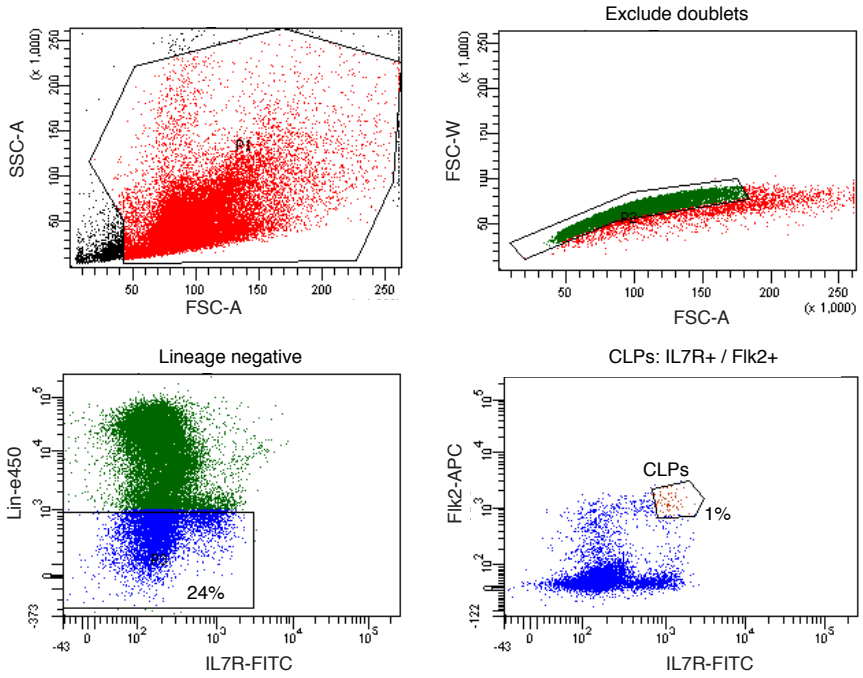


Figure S11. Predicting transcription factor mediated enhancer regulation during hematopoiesis by a logistic regression model. **(A)** Heatmap showing t-statistics for the coefficient estimates for a generalized linear regression for each factor (rows) in different cell type (columns). **(B)** Contour lines showing the density of H3K4me1 vs. logistic model predictions for each enhancer X cell-type. The x-axis shows the enhancer H3K4me1 levels, and the y-axis the logistic classifier score. **(C)** Hierarchy in Cebp factors in the myeloid lineage development. **Top** Heatmap showing p-values of transcription factor motif K-S test for Cebpa, Cebpb, and Cebpe in the indicated cell types. White indicates either no significant enrichment or no RNA expression (Method). **Bottom** Bar plots showing RNA expression of Cebpa, Cebpb and Cebpe in the indicated cell types.

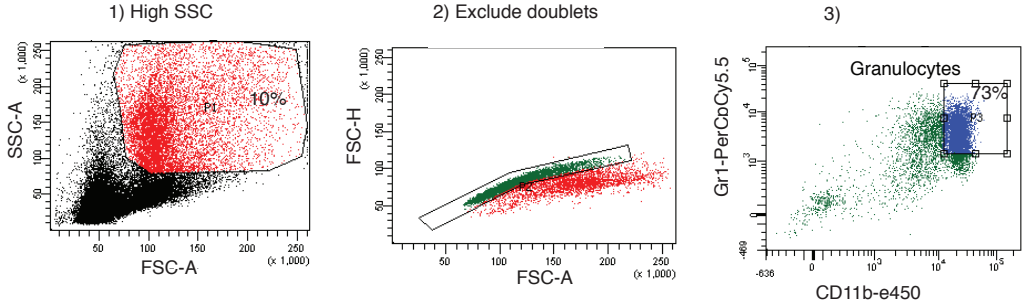
Progenitors



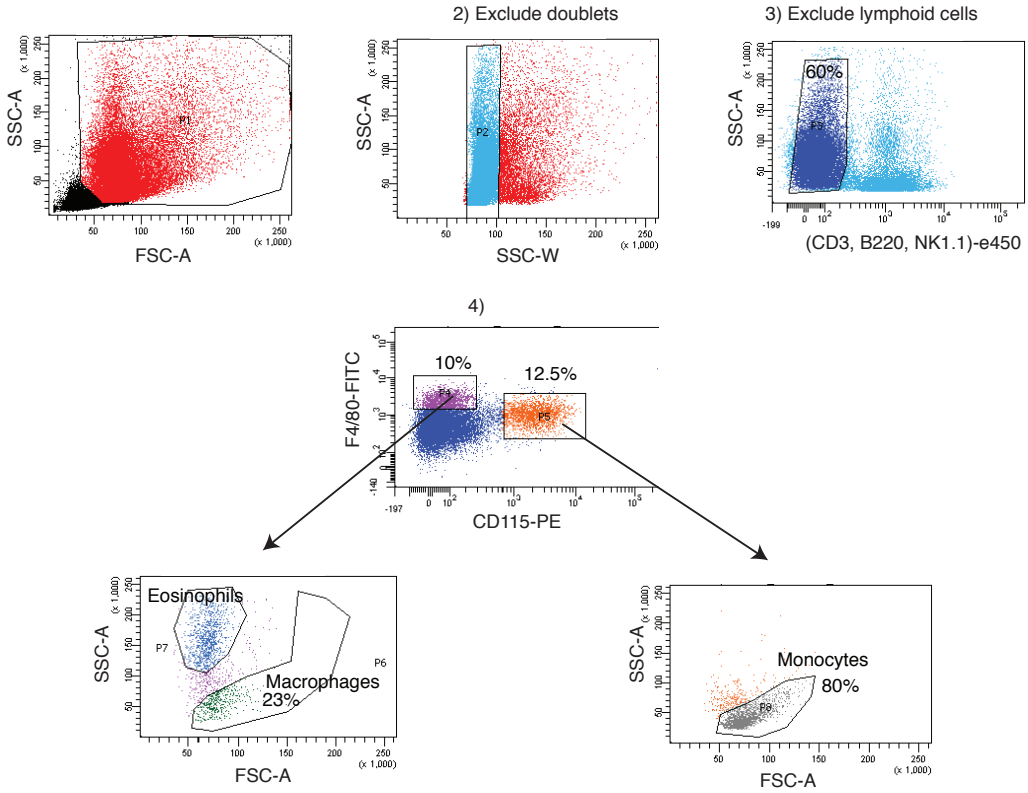
CLPs


SSC-A
FSC-A
R1
Exclude doublets
FSC-W
FSC-A
Lineage negative
Lin-e450
IL7R-FITC
24%
CLPs: IL7R+ / Flk2+
Flk2-APC
IL7R-FITC
CLPs
1%

Granulocytes

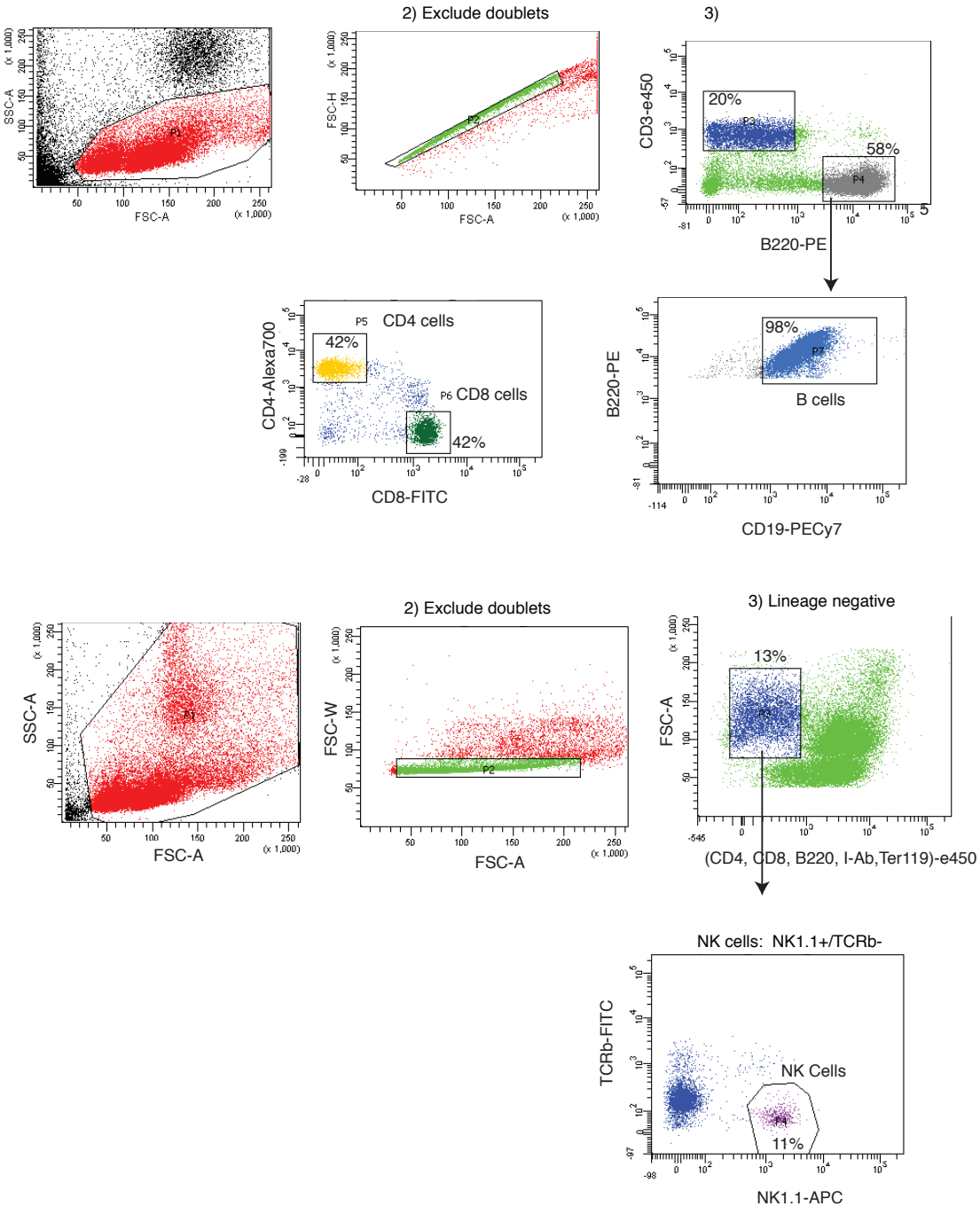


Monocytes and Macrophages



Lara, Weiner *et al.*

Lymphoid Cells



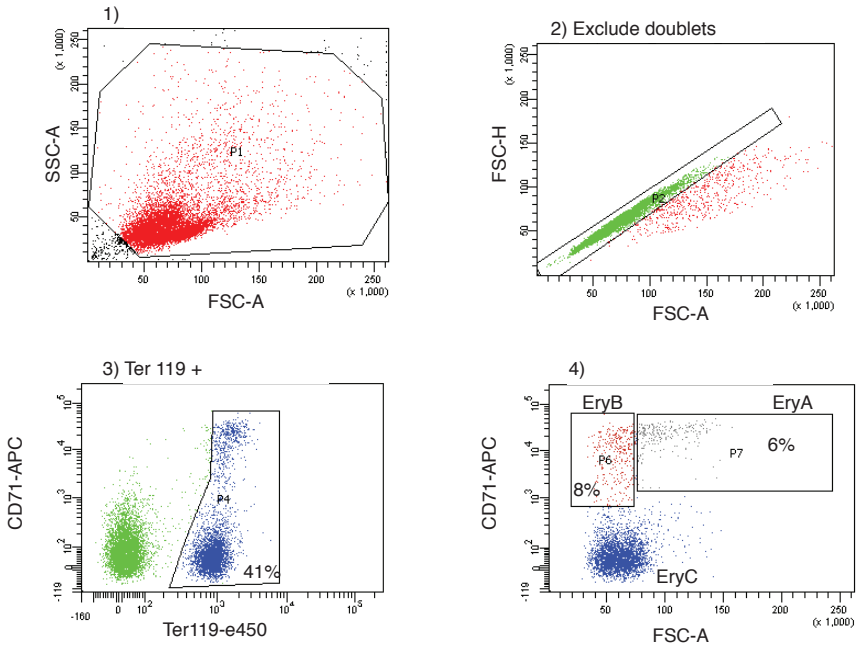


Figure S12. FACS plots showing the strategy to purify the 16 cell populations analyzed in this study. (A) Progenitor populations (except CLPs). **(B)** Common Lymphoid Progenitors (CLPs). **(C)** Mature myeloid cell populations. **(D)** Mature lymphoid cell populations. (E) Mature erythroid cell populations.

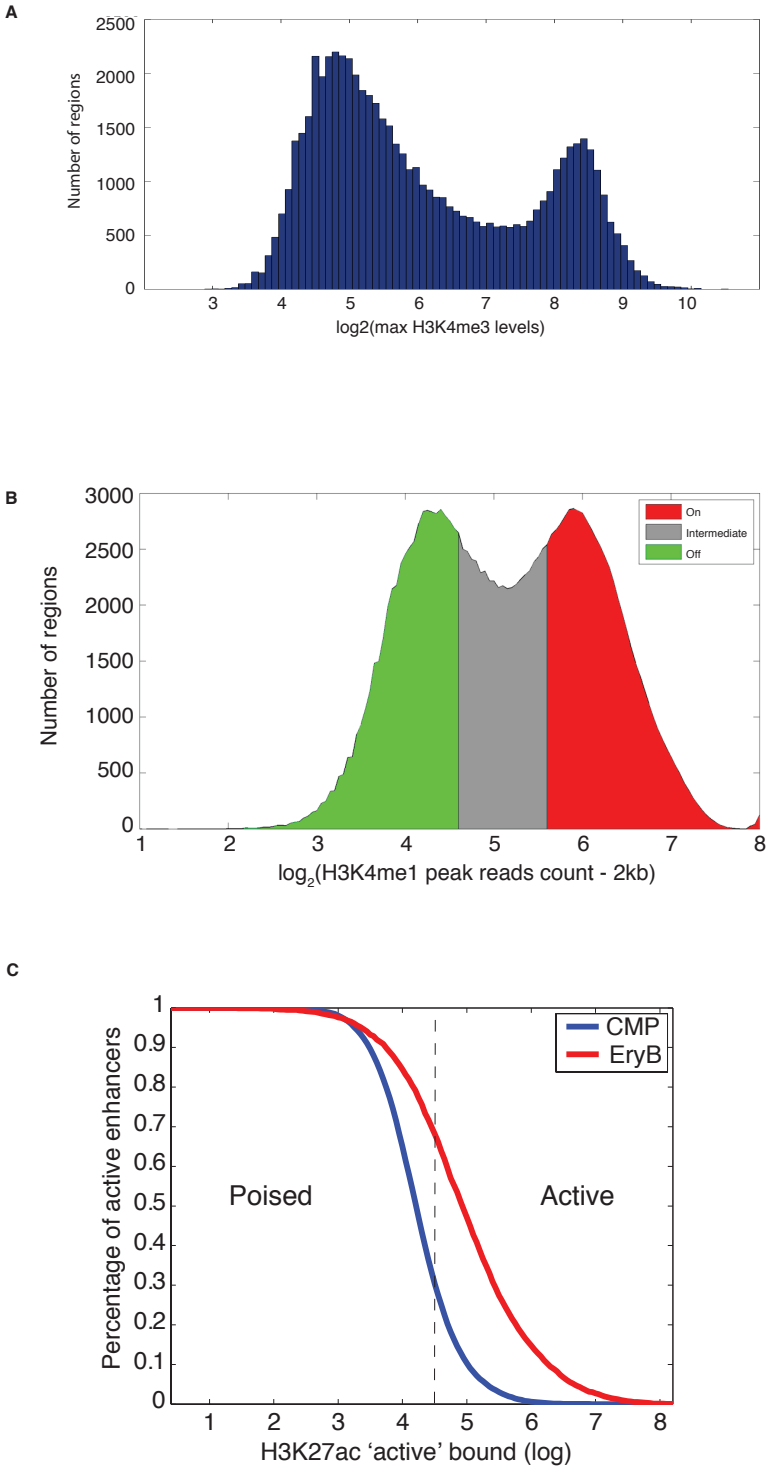


Figure S13. (A) Histogram of H3K4me3 read counts (in log scale) over H3K4me1 peaks. **(B)** Histogram of H3K4me1 peaks read coverage in all 16 cell types together, green mark 'off' enhancers, red for 'on' enhancers and gray for 'intermediate. **(F)** Percentage of 'active' enhancers in CMP (blue) and EryB (red) as a function of H3K27ac threshold, dash line represent the threshold used to separate 'active' and 'poised' enhancers.

Table S1. Hematopoietic enhancers catalog (excel)

Table S2. RNA-seq expression data (excel)

Table S3. Primers Table

PRIMER NAME	SEQUENCE AND MODIFICATIONS
Universal Adaptor	5′-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3′, where * indicates phosphothionate modification.
Indexed Adaptors	5'-GATCGGAAGAGCACACGTCTGAACTCCAGTCACXXXXXXATCTCGTATGCCGTCTTCTGTT- 3', where XXXXXX is the barcode for sample multiplexing
Y- Shaped Indexed adaptors	- Anneal ssUniversal Adaptor and ssIndexed Adaptors to obtain Y-shaped Indexed Adaptors.
PCR_forward	5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC-3′
PCR_Reverse	5′-CAAGCAGAAGACGGCATACGAGAT-3′
Barcoded RT primer	5'-CGATTGAGGCCGGTAATACGACTCACTATAGGGGCGACGTGTGCTCTTCCGATCTXXXXXXNNNNTTTTTTTTTTTTTTTTTTTTN-3', where XXXXXX is the cell barcode and NNNN is the RMT
Ligation adapter	5'-AGATCGGAAGAGCGTCGTGTAG-3', modified with a phosphate group at 5' and a C3 spacer (blocker) at the 3'
Second RT primer	5'-TCTAGCCTTCTCGCAGCACATC-3'
P5_Rd1 PCR forward	5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'
P7_Rd2 PCR reverse	5'-CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'

Table S4. FACS Antibodies

Marker	Fluorophore	Clone	
B220	e450	RA3-6B2	eBioscience
B220	PE	RA-6B2	eBioscience
c-kit	APC	2B8	eBioscience
CD115	PE	AFS98	eBioscience
CD11b	e450	M1/70	eBioscience
CD11b	APC	M1/70	eBioscience
CD19	PE-Cy7	Bio1D3	eBioscience
CD3	e450	17A2	eBioscience
CD34	FITC	RAM34	eBioscience
CD4	e450	GK1.5	eBioscience
CD4	Alexa700	GK1.5	eBioscience
CD71	APC	R17217	eBioscience
CD8	FITC	53-6.7	eBioscience
CD8a	e450	53-6.7	eBioscience
F4/80	FITC	BM8	eBioscience
FcgR-II	PE-Cy7	93	eBioscience
Flk2	PE	A2F10	eBioscience
Gr1	e450	RB6-8C5	eBioscience
Gr1	PerCP-Cy5.5	RB6-8C5	Biolegend
I-Ab	PacBlue	AF6-120.1	eBioscience
IL7-R	FITC	A7R34	eBioscience
NK1.1	e450	PK136	eBioscience
NK1.1	APC	PK136	eBioscience
Sca-1	PerCP-Cy5.5	D7	eBioscience
TCR-β	FITC	H57-597	eBioscience
Ter119	e450	TER-119	eBioscience

References

1. L. Ho, G. R. Crabtree, Chromatin remodelling during development. *Nature* **463**, 474–484 (2010). Medline doi:10.1038/nature08911

2. V. W. Zhou, A. Goren, B. E. Bernstein, Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* **12**, 7–18 (2011). Medline doi:10.1038/nrg2905

3. S. Ghisletti, I. Barozzi, F. Mietton, S. Polletti, F. De Santa, E. Venturini, L. Gregory, L. Lonie, A. Chew, C. L. Wei, J. Ragoussis, G. Natoli, Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32**, 317–328 (2010). Medline doi:10.1016/j.immuni.2010.02.008

4. E. M. Mercer, Y. C. Lin, C. Benner, S. Jhunjhunwala, J. Dutkowski, M. Flores, M. Sigvardsson, T. Ideker, C. K. Glass, C. Murre, Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity* **35**, 413–425 (2011). Medline doi:10.1016/j.immuni.2011.06.013

5. G. Wei, L. Wei, J. Zhu, C. Zang, J. Hu-Li, Z. Yao, K. Cui, Y. Kanno, T. Y. Roh, W. T. Watford, D. E. Schones, W. Peng, H. W. Sun, W. E. Paul, J. J. O'Shea, K. Zhao, Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating $CD4^+$ T cells. *Immunity* **30**, 155–167 (2009). Medline doi:10.1016/j.immuni.2008.12.009

6. R. Revilla-i-Domingo, I. Bilic, B. Vilagos, H. Tagoh, A. Ebert, I. M. Tamir, L. Smeenk, J. Trupke, A. Sommer, M. Jaritz, M. Busslinger, The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J.* **31**, 3130–3146 (2012). Medline doi:10.1038/emboj.2012.155

7. T. S. Mikkelsen, M. Ku, D. B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T. K. Kim, R. P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E. S. Lander, B. E. Bernstein, Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007). Medline doi:10.1038/nature06008

8. J. Seita, I. L. Weissman, Hematopoietic stem cell: Self-renewal versus differentiation. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2**, 640–653 (2010). Medline doi:10.1002/wsbm.86

9. S. Doulatov, F. Notta, E. Laurenti, J. E. Dick, Hematopoiesis: A human perspective. *Cell Stem Cell* **10**, 120–136 (2012). Medline doi:10.1016/j.stem.2012.01.006

10. A. H. Shih, O. Abdel-Wahab, J. P. Patel, R. L. Levine, The role of mutations in epigenetic regulators in myeloid malignancies. *Nat. Rev. Cancer* **12**, 599–612 (2012). Medline doi:10.1038/nrc3343

11. S. Heinz, C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, C. K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010). Medline doi:10.1016/j.molcel.2010.05.004

12. H. Ji, L. I. Ehrlich, J. Seita, P. Murakami, A. Doi, P. Lindau, H. Lee, M. J. Aryee, R. A. Irizarry, K. Kim, D. J. Rossi, M. A. Inlay, T. Serwold, H. Karsunky, L. Ho, G. Q. Daley, I. L. Weissman, A. P. Feinberg, Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467**, 338–342 (2010). Medline doi:10.1038/nature09367

13. T. S. Furey, ChIP-seq and beyond: New and improved methodologies to detect and characterize protein-DNA interactions. *Nat. Rev. Genet.* **13**, 840–852 (2012). Medline doi:10.1038/nrg3306

14. P. Shankaranarayanan, M. A. Mendoza-Parra, M. Walia, L. Wang, N. Li, L. M. Trindade, H. Gronemeyer, Single-tube linear DNA amplification (LinDA) for robust ChIP-seq. *Nat. Methods* **8**, 565–567 (2011). Medline doi:10.1038/nmeth.1626

15. M. Garber, N. Yosef, A. Goren, R. Raychowdhury, A. Thielke, M. Guttman, J. Robinson, B. Minie, N. Chevrier, Z. Itzhaki, R. Blecher-Gonen, C. Bornstein, D. Amann-Zalcenstein, A. Weiner, D. Friedrich, J. Meldrim, O. Ram, C. Cheng, A. Gnirke, S. Fisher, N. Friedman, B. Wong, B. E. Bernstein, C. Nusbaum, N. Hacohen, A. Regev, I. Amit, A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* **47**, 810–822 (2012). Medline doi:10.1016/j.molcel.2012.07.030

16. Materials and methods are available as supplementary materials on *Science* Online.

17. C. M. Rivera, B. Ren, Mapping human epigenomes. *Cell* **155**, 39–55 (2013). Medline doi:10.1016/j.cell.2013.09.011

18. M. U. Kaikkonen, N. J. Spann, S. Heinz, C. E. Romanoski, K. A. Allison, J. D. Stender, H. B. Chun, D. F. Tough, R. K. Prinjha, C. Benner, C. K. Glass, Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell* **51**, 310–325 (2013). Medline doi:10.1016/j.molcel.2013.07.010

19. A. Rada-Iglesias, R. Bajpai, T. Swigut, S. A. Brugmann, R. A. Flynn, J. Wysocka, A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011). Medline doi:10.1038/nature09692

20. M. P. Creyghton, A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, R. Jaenisch, Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21931–21936 (2010). Medline doi:10.1073/pnas.1016071107

21. T. A. Wynn, A. Chawla, J. W. Pollard, Macrophage biology in development, homeostasis and disease. *Nature* **496**, 445–455 (2013). Medline doi:10.1038/nature12034

22. R. E. Thurman, E. Rynes, R. Humbert, J. Vierstra, M. T. Maurano, E. Haugen, N. C. Sheffield, A. B. Stergachis, H. Wang, B. Vernot, K. Garg, S. John, R. Sandstrom, D. Bates, L. Boatman, T. K. Canfield, M. Diegel, D. Dunn, A. K. Ebersol, T. Frum, E. Giste, A. K. Johnson, E. M. Johnson, T. Kutyavin, B. Lajoie, B. K. Lee, K. Lee, D. London, D. Lotakis, S. Neph, F. Neri, E. D. Nguyen, H. Qu, A. P. Reynolds, V. Roach, A. Safi, M. E. Sanchez, A. Sanyal, A. Shafer, J. M. Simon, L. Song, S. Vong, M. Weaver, Y. Yan, Z. Zhang, Z. Zhang, B. Lenhard, M. Tewari, M. O. Dorschner, R. S. Hansen, P. A. Navas, G. Stamatoyannopoulos, V. R. Iyer, J. D. Lieb, S. R. Sunyaev, J. M. Akey, P. J. Sabo, R. Kaul, T. S. Furey, J. Dekker, G. E. Crawford, J. A. Stamatoyannopoulos, The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012). Medline doi:10.1038/nature11232

23. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang, W. J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013). Medline doi:10.1038/nmeth.2688

24. L. A. Cirillo, F. R. Lin, I. Cuesta, D. Friedman, M. Jarnik, K. S. Zaret, Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**, 279–289 (2002). Medline doi:10.1016/S1097-2765(02)00459-8

25. A. B. Cantor, S. H. Orkin, Transcriptional regulation of erythropoiesis: An affair involving multiple partners. *Oncogene* **21**, 3368–3376 (2002). Medline doi:10.1038/sj.onc.1205326

26. W. Ouyang, M. O. Li, Foxo: In command of T lymphocyte homeostasis and tolerance. *Trends Immunol.* **32**, 26–33 (2011). Medline doi:10.1016/j.it.2010.10.005

27. T. Matsuyama, T. Kimura, M. Kitagawa, K. Pfeffer, T. Kawakami, N. Watanabe, T. M. Kündig, R. Amakawa, K. Kishihara, A. Wakeham, Targeted disruption of IRF-1 or IRF-2 results in abnormal type I IFN gene induction and aberrant lymphocyte development. *Cell* **75**, 83–97 (1993). Medline doi:10.1016/0092-8674(93)90681-F

28. R. Yamanaka, C. Barlow, J. Lekstrom-Himes, L. H. Castilla, P. P. Liu, M. Eckhaus, T. Decker, A. Wynshaw-Boris, K. G. Xanthopoulos, Impaired granulopoiesis, myelodysplasia, and early lethality in CCAAT/enhancer binding protein epsilon-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13187–13192 (1997). Medline doi:10.1073/pnas.94.24.13187

29. D. A. Jaitin, E. Kenigsberg, H. Keren-Shaul, N. Elefant, F. Paul, I. Zaretsky, A. Mildner, N. Cohen, S. Jung, A. Tanay, I. Amit, Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014). Medline doi:10.1126/science.1247651

30. P. McCullagh, J. Nelder, *Generalized Linear Models*, *Second Edition* (Chapman & Hall, ed. 2, 1989).